

Elsevier Editorial System(tm) for Journal of
Computational Science
Manuscript Draft

Manuscript Number: JOCS-D-16-00260R1

Title: How People Learn while playing Serious Games: A Computational
Modelling Approach

Article Type: Full Length Article

Keywords: serious gaming; learning; simulation; flow theory; methodology

Corresponding Author: Professor Wim Westera,

Corresponding Author's Institution: Welten Institute

First Author: Wim Westera

Order of Authors: Wim Westera

Abstract: This paper proposes a computational modelling approach for investigating the interplay of learning and playing in serious games. A formal model is introduced that allows for studying the details of playing a serious game under diverse conditions. The dynamics of player action and motivation is based on cognitive flow theory, which is expressed in quantitative terms for this purpose. Seven extensive simulation studies involving over 100,000 iterations have demonstrated the stability of the model and its potential as a research instrument for serious gaming. The model allows researchers to deeply investigate quantitative dependences between relevant game variables, gain deeper understanding of how people learn from games, and develop approaches to improving serious game design.

COVER LETTER

Heerlen, November 18th 2016

Dear editor of the Journal of Computational Science

Herewith the re-submission **JOCS-D-16-00260R1**. I appreciate the reviewers' comments and have revised the paper accordingly.

Again I declare that the manuscript has not been submitted, considered or published elsewhere. I have closely followed your style and submission instructions (e.g. anonymised version). As the only author of the manuscript, I will obviously be the corresponding author (see details below).

Looking forward to your response,

kindest regards

Wim Westera

Wim Westera

Welten Institute, Research Centre for Learning, Teaching and Technology

Faculty of Psychology and Educational Sciences,

Open University of the Netherlands

Valkenburgerweg 177

6419 AT Heerlen

The Netherlands

Telephone: +31-45-5762408

Wim.westera@ou.nl , Wim.westera@gmail.com

Skype: celstec-wim.westera

www.open.ou.nl/wim

www.thedigitalturn.co.uk

Wim Westera is full a professor in learning media, specialised in media for education, in particular gaming and simulation. He holds a PhD in physics and mathematics, and has worked in educational media development and educational technology since the 1980s. He is leading RAGE, which is the principal H2020 research project on applied gaming (rageproject.eu).



How People Learn while playing Serious Games: A Computational Modelling Approach

1. Introduction

1.1.The absorbing nature of serious games

Driven by the successes of the leisure game industry, games increasingly find their way into non-leisure contexts, serving serious purposes. These so-called “serious games” span a wide range of application areas, including training and learning, awareness raising and sensitisation, as well as marketing and the advancement of cultural engagement [1,2]. This paper focuses particularly on games for learning. A principal argument for using games in education and training is the engaging nature of gaming and the motivational power that games display: the ability of hooking and absorbing players in such a way that they can hardly stop playing [3-5]. This potential is ascribed to their dynamic, responsive and visualised nature, which goes along with novelty, variation and choice, effecting strong user involvement and providing penetrating learning experiences [4]. In addition, serious games allow for safe experimentation in realistic environments, stimulate problem ownership by role adoption, and allow for learning-by-doing approaches, which support the acquisition of tacit and contextualised knowledge [6].

1.2.The inherent complexity of games

Games are inherently complex constructs comprising knotty structures of highly interrelated components that may vary over time. Björk and Holopainen [7] qualify game design and development as a semi-formalised, fuzzy and incoherent domain, which eclectically combines various approaches that cannot be fully covered by prescriptive or even descriptive theories. For serious games, which pair game design with instructional design, the complexity may even be larger because of the multiplication of two ill-structured domains, requiring the

cautious balancing of “playful” game mechanics and “serious” instructional principles [8,9]. To some extent games suffer from an impenetrable interior. Salen and Zimmerman [10] note that the link between the designed structural properties of a game and the effected user experience remains often unclear, because of the vast space of game states and the large number of trajectories a player could travel through the game’s state space. Consequently, different players may have different game experiences as they engage in different trajectories and game events and thereby experience different cumulative narratives: different runs of a game may be very different. It would not be sufficient to test a game for the “average pathway”, because no single player would ever traverse the “average pathway”. In serious games it may be hard to tell how individual decisions will impact on the game experience and how this affects the game’s effectiveness for learning [11].

1.3. A computational modelling approach

This paper presents a computational model for simulating how people learn while playing serious games. Its main purpose is to allow and support researching what happens during playing a game under a variety of conditions. Such model should reflect the structural dynamics of the game and should help to enlighten its capacity as a learning aid. The research is decomposed into the following research questions:

- How to formulate an expressive computational model for the process of playing a serious game that avoids inherent complexity?
- To what extent is such computational model capable of producing stable results?
- To what extent does the model produces acceptable outcomes, which are consistent with empiricism on learning from games?

The paper is setup as follows. First the main methodological considerations for this study will be summarised. Second, the model’s starting points and its grounding in theory will be made explicit. Third, the computational model will be defined and substantiated. Fourth, a variety

of simulation experiments will be reported, which include over 100,000 simulation runs in total. The paper is concluded by discussing the results and their potential implications.

2. Methodological considerations

2.1.Methods fight

For many years the social sciences have shown a strife between methodological camps with on the one hand the empirical hypothesis testing framework, which tries to validate hypotheses by subjecting observed measures to statistical analysis, and on the other hand deductive modelling (e.g. game theoretical methods), which aims to specify benefits and costs schemes for explaining individuals' behaviours [12]. The empirical methods of experimentation have been persistently criticised for their unintelligent data crunching, limited explanatory power, their biased focus on positive effects and their arbitrary significance measures [13,14]. Game theorists in turn, who describe gaming in terms of strategic decision making by rational human players, have been blamed for their believe that formal theory doesn't require empirical referents [12]. Less well known as an additional research method in social sciences is computational systems modelling, which incorporates aspects of both empirical research and game-theoretical research approaches by capturing the individuals' behaviours in behavioural rules along with a set of contextual parameters and constraints, and produce a dynamic model that recreates observed phenomena [12]. Although computational system modelling has been criticised for allowing large parameter spaces, which easily lead to model overfitting, in the last decades computational methods have been successfully applied in diverse complex domains, ranging from atomic scale protein design and nuclear fusion to superconductivity and a billion-particles simulation of the Milky Way. Various authors advocate the widespread application of computational models for the integration of theoretical, technical and empirical research [15].

2.2.Serious gaming as an emerging field of research

In the domain of serious games the contributions from computer science have gained importance, particularly because of the impact of advanced digital game technologies [16]. Still, most research adheres to empirical research particularly grounded in the learning sciences. These aim to professionalise teaching on the basis of sound, empirically supported instructional methods rather than viewing teaching as an art, driven by intuition and feeling [17,18]. The link with gaming is readily in experiential learning, learning-by-doing, motivation theory, multimedia learning, social and collaborative learning, connectionism or networked learning. But those theories are largely qualitative and descriptive by nature and seem to lack the level of formalisation and precision required for making valid predictions. Because of this, research in the learning sciences as well as the instruments used have been persistently criticised [15,19-22]. Despite the valuable insights and confirmations that learning sciences research has produced over the last decades, it has not been capable of making predictions about instructional situations.

2.3.Strengthening multidisciplinary research

Given the multidisciplinary nature of the field of serious gaming, its research would require a close connection between its constituting domains such as learning sciences, game theory and computer sciences. However, the cross-fertilisation between these sub-domains has been weak, not just because of different cultures and paradigms in these disciplines, but also because of the disparate backgrounds and expertise that is required. Apparently this is the inevitable fate of any emerging multidisciplinary field. Current research on serious games is dominated by case studies, that is, the research focuses on case-by-case descriptions of a highly qualitative nature about particular games under study and its appreciations by users. Although an increasing body of evidence is becoming available that reveals the effectiveness of serious games for learning, various authors note that many studies fail to evaluate the

educational effectiveness of serious games in a rigorous manner and they call for quantitative research and comprehensive frameworks for increased scientific robustness [23,24]. Still most studies focus on post-practice results and they neglect what actually happens during playing games. Given the inherent complexity of serious games, representing a game by a computational model would allow for testing and evaluating a wide variety of behaviours and thus would allow for a more representative view on game experiences. Once computational models have been generalised and verified for explaining behavioural phenomena, e.g. playing and learning in a serious game, the model could – in principle - be run and rerun to reveal behavioural diversity across different personal traits and external conditions.

3. Model starting points and ingredients

Before elaborating the serious gaming model, first the main issues and starting points will be reported.

3.1.Avoiding the combinatorial explosion of game states and player states

The deterministic idea that knowing all potential game states and all player states and their progression over time would eventually allow us to devise the player's optimal learning strategy, that is, the optimal trajectory through the game state space, is illusive. Game representation is likely to suffer from a combinatorial explosion of game states. Even a simple game such as tic-tac-toe (noughts and crosses) has a state space up to $3^9=19,683$ different states (neglecting any symmetries) allowing for $9!=362,880$ different trajectories. Taking into account symmetries and including games that end within 9 moves only, the number of trajectories is still 26,830 [25], an inconceivable number way too high to even be depicted in a game tree.

Likewise, the full user profile of a player includes a wide variety of factors that could be of relevance for learning, e.g. intelligence [26,27], motivation [28-31], knowledge and skills [32-

35], personality traits [36], emotional states [37], playing styles [38], learning styles [39-41], learning dispositions [42]. However, most of these factors cannot be tracked during the process of play (or even outside), because of either their qualitative nature or the lack of reliable measurement methods. Moreover, there is no coherent theory available that would connect these variables into an overall causal framework for making predictions.

For being successful at devising a practicable computational model of serious gaming simplification is essential. The model should avoid the detailed game states and micro-level decisions in the game, while instead it should focus on meso-level aggregates that constitute meaningful activities, directly originating from the game scenario, e.g. write a note, interview a person, track a hidden object, buy supplies, navigate. Rather than representing the game by a trajectory through its game's full parameter state space the game will be described by the successive moves across such meso-level game activities. Playing the game can thus be understood as a trajectory along a limited number of sizable game activities, which are the meaningful pursuits within the context of the game and fit the player's strategies and progression toward the game's objectives. Also it is in agreement with the observation that from the player's perspective most games are quite surveyable, since they perceive the game as a series of successive tasks they accomplish. Tic-tac-toe with its nine player steps would be a point in case.

3.2. Grounding on principal concepts of learning and instructional design

For being effective as a learning tool serious games have to offer appropriate conditions and activities that contribute to achieving pursued learning objectives. The model will incorporate a number of essential factors identified in learning psychology and instructional design research. First, it is acknowledged that activities in the serious game environment may differ in complexity and cognitive demand. Also some activities may be more attractive and pleasant than others. The effectiveness of learning when passing through a game activity is

1 readily influenced by these factors. Also player characteristics may change the effectiveness
2 of learning, for instance the players' motivations or deficiencies in prior knowledge required
3
4 for a new task. According to Gee [5] tying such factors together is the secret of serious
5
6 gaming, which is not in the high quality graphics, but in the underlying architecture, which
7
8 balances the challenges offered to the player with the players' abilities seeking at every point
9
10 to be hard enough to be just doable. In psychological terms this mechanism is easily linked
11
12 with Csikszentmihalyi's theory of cognitive flow [43] and Vygotsky's zone of proximal
13
14 development [44]. Cognitive flow refers to a mental state characterised by extreme
15
16 involvement, concentration, engrossment, restricted awareness, altered sense of time,
17
18 insensitiveness to hunger and insensitiveness to fatigue [43]. Such state of intensive mental
19
20 activity is highly favourable for sustained learning. Achieving cognitive flow requires a
21
22 cautious adjustment of the challenges offered to the player's abilities. If the challenges are too
23
24 hard the player is likely to become frustrated and to lapse into apathy. In contrast, if the
25
26 challenges are too easy, the player is like to get bored and lose interest. Under the right
27
28 balance players may be pulled into cognitive flow and benefit from the intensified
29
30 concentration and involvement. In addition, Vygotsky's theory of social development
31
32 suggests that learners should be challenged slightly beyond the boundaries of their abilities,
33
34 while avoiding both frustration and boredom. This stretching of the boundaries is preserving
35
36 the curiosity and engagement that are needed for learning new things that go beyond existing
37
38 knowledge and routines. A well-designed serious game would incorporate an appropriate
39
40 mechanism for balancing challenges and player's abilities. The proposed computational
41
42 model will include the cognitive flow mechanism.
43
44
45
46
47
48
49
50
51
52
53

54 ***3.3. System components***

55 Any model that covers the process of playing a serious game should rely on representations of
56
57 the following sub-systems: the knowledge model, the game model, the player model, which
58
59
60
61
62
63
64
65

all come together in the frame of operation, that is, the overall process of playing the game (Figure 1).

<<< FIGURE 1 ABOUT HERE>>>

Figure 1. Serious gaming model components.

The knowledge model defines the learning objectives to be attained. The game model characterises the game by reflecting the game mechanics, game challenges and game content. The player model represents the characteristics and mental states of the player during the process of play. The frame of operation refers to the productive interactions between the player and the game. For simplifying the overall model, each of these subsystems should be described at a sufficiently high level, while preserving the most relevant factors. Frugality with respect to the wide range of variables that are available for describing the processes and conditions of serious gaming is also dictated to avoid overfitting of the model. Once a simple model would have proven its usefulness, it could be further detailed.

3.4.The knowledge model

This subsystem reflects the stable structure of the knowledge to be learned in the game. For practical reasons this paper does not distinguish between knowledge, skills and competences, but instead uses the term knowledge as a transcending, inclusive concept indicating the things to be learned. As both in education and training the knowledge model serves as the benchmark for assessment and certification, knowledge is often expressed as learning objectives [32,33,45,46]. Generally, the knowledge model can be represented as a hierarchical framework of interrelated knowledge elements, while child nodes in the hierarchy have a precedence relationship with their parent nodes. This means that for being able to learn new knowledge from a parent node the (partial) mastery of the child nodes' knowledge is presupposed. The knowledge model is static by its nature of expressing the benchmark of

required learning outcomes. In addition several instances of this hierarchical structure can be used to represent 1) the baseline prior knowledge requirements for being able to play the game, 2) the player's prior knowledge state, 3) the player's progressing knowledge state during the process of play, and 4) the observable behaviours and performances that provide evidence for the mastery of underlying knowledge, the evidence model [45].

3.5.The game model

Rather than describing games by their numerous potential game states, which would account for every detailed player interaction (e.g. mouse clicks), a game is considered as an environment that offers the players a coherent set of activities, e.g. challenges, tasks, assignments, missions, scenes or levels that need to be passed through. Thereby a game activity is conceived as a higher level aggregate of micro-actions, that constitute a well-defined chunk of the game scenario: it should offer a clear challenge, it has a well-delineated scope, and its completion goes with a clear result or achievement defined by a completion criterion. Hence, a game activity transcends the level of elementary user actions (which are typically button clicks or keyboard strokes). It does describe only a part of the game rather than the whole. Thereby a game activity is very similar to one or more learning tasks or learning activities in a lesson. Thus, each game activity offers a learning experience that contributes to some of the expected learning outcomes, which in principle refer to the elements of the knowledge model. Given the set of game activities, a serious game is represented as a network of activity nodes. Playing a serious game can be interpreted as doing a trajectory through the network of activities. Thereby it is conceptually related to the Travelling Salesperson Problem, which is stated as finding the shortest route along a number nodes (locations) in a network [47]. The problem is known as an NP-hard problem in combinatorial optimisation, but for a limited number of nodes practical solutions can be found by brute force. The conditional dependences that are reflected in the hierarchic knowledge

1 model will generally translate into sequential dependences between the nodes in the game
2 activities network, which means that not all trajectories through the activity network will be
3 practical, given the required prior knowledge for each node. In addition, the sequence order of
4 activities may be restricted by the game narrative, which imposes a logical or causal order of
5 game events. Because of these constraints the number of acceptable trajectories through the
6 activity network will be appreciably reduced.
7
8
9
10
11
12

13
14
15
16 A game activity should not only be described by the knowledge elements that it is supposed to
17 cover, but it also needs an indicator for its complexity as to allow for determining whether or
18 not the activity matches the conditions for the player's cognitive flow: the next step in the
19 trajectory should not be too complex as to avoid frustration, it should not be too simple as to
20 avoid boredom. In addition, not all activities will offer the same depth of the learning
21 experience, that is, the engagement induced by the activities may greatly differ. Deep and
22 engaging experiences will be likely to better learning outcomes. This engagement dimension
23 should basically cover the inherent potential of the game activity to engage and immerse the
24 player in the game: it reflects the attractiveness of the game (including game mechanics,
25 narrative and style) offered. In sum, the main attributes of a serious game activity are the
26 knowledge elements that it covers, the complexity of the activity and the attractiveness of the
27 activity.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

47 ***3.6.The player model***

48
49 As the goal of serious gaming is to provide an effective means for training and learning, the
50 player's knowledge state obviously is a key variable. While moving from one game activity to
51 the other the players gradually extend their knowledge with new knowledge. A simplified
52 player model should limit itself to only include primary factors, while possibly neglecting the
53 underlying causal variables. Basically the learning achievements resulting from a learning
54
55
56
57
58
59
60
61
62
63
64
65

1 activity would depend on the players' prior knowledge on the one hand, and the players'
2 engagement and motivation to learn on the other hand. Meeting prior knowledge requirements
3 is an important factor for the successful completion of an activity [48]. Players suffering
4 substantial knowledge deficiencies will have a hard time to learn new things that are out of
5 reach. Likewise the players' motivation is an essential factor for productive learning, as a high
6 motivation will go with increased engagement, attention, concentration and intensity [28-31].
7 Whilst innate factors such as intelligence and personality tend to be stable, motivation may
8 change during the game as a result of the interactions performed. While the prior knowledge
9 refers to the learners' knowledge states, motivation refers to their personal attitudes and
10 ambitions. These are exactly the key dimensions that reflect the potential of serious games:
11 learning new knowledge, while benefitting from the motivational power of games.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 ***3.7.The frame of operation***

29 Given the knowledge model, the game model and the player model, the frame of operation
30 reflects the process of play, which describes the player's interactions, navigation and history
31 of activities and achievements. The process of play is the effect of the players' goal-oriented
32 behaviours under the constraints imposed by the game environment and the players' personal
33 capabilities. The dynamics of the proposed model will be based on a quantified version the
34 cognitive flow theory, which cautiously takes into account the balancing between the player's
35 capabilities and the challenges offered. In sum, the model starts from cognitive flow theory
36 and takes into account the complexity and attractiveness of game activities, the knowledge
37 objectives that are addressed by each activity, the player's progressive knowledge state for
38 each node of the knowledge tree, the required prior knowledge for each game activity, the
39 level of challenge offered, the player's motivation, the player's overall intelligence and the
40 resulting effectiveness of the learning.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

4. Model elaboration

4.1. Defining the knowledge model

This model is represented as a so-called k-ary tree of learning objectives (e.g. knowledge, skills, competences). Precedence relationships between nodes at different levels in the tree are represented as parent-child edges. A parent node may refer to multiple child nodes. Note that in educational sciences and cognitive psychology such knowledge or skills models are often referred to as hierarchies rather than trees [32,33]. This paper, however, will comply with graph terminology and refer to knowledge trees rather than hierarchies. For reasons of simplification the knowledge model is described as a perfect k-ary tree, where k is called the degree, which is the number of child nodes of each parent node. Assuming a perfect k-ary tree, however, does not imply a fundamental constraint, as it would still allow for transforming it into any irregular tree model by removing selected nodes and edges from the tree.

For describing the k-ary knowledge tree, the following notation is used:

N_g The number of levels in the knowledge tree

N_c The number of children each parent node has (the tree's degree k)

N_n The number of nodes in the knowledge tree

It can be shown that the total number of nodes is given by

$$N_n = \sum_{i=1}^{N_g} N_c^{(i-1)} \quad (1)$$

4.2. Defining the game model

As the serious game is supposed to support the mastery of the objectives described by the knowledge model, all knowledge elements (nodes in the hierarchy) are connected with game activities. Each game activity may address multiple nodes of the knowledge tree. At the same time multiple activities may address the same knowledge node. Although, in principle,

the mapping relationships between game activities and knowledge nodes will be diverse, we assume for reasons of simplicity that each game activity covers a fixed number of knowledge nodes, and that the number of occurrences of the knowledge nodes is likewise a fixed number.

This can be expressed as follows:

N_a The number of activities that make up the game

N_i The incidence of each knowledge node in the game

N_k The number of knowledge nodes that are supported by an activity

This simplification does not pose a basic restriction of the proposed computational model as it only affects the model's input data.

As the mapping between knowledge nodes and the game activities in this simple case is unambiguous, it can be shown that the number of game activities is given by

$$N_a = \frac{N_n \cdot N_i}{N_k} \quad (2)$$

Each game activity j will have the following attributes:

$K(i,j)$ This indicates the mapping of knowledge nodes i to game activity j, while each game activity j addresses N_m knowledge elements

$C(j)$ The complexity assigned to game activity j (indicated as a fraction)

$A(j)$ The attractiveness of game activity j (indicated as a fraction)

4.3. Defining the player model

4.3.1. The mastery level

First, during the game the player's progress should be tracked for each node of the knowledge tree. The player's knowledge model would be an instance of the general knowledge tree, be it annotated with the player's level of mastery for each knowledge node. Passing through a game activity would contribute to the mastery of the associated knowledge

nodes, be it that in general the mastery will not be perfect. Inevitably each player is cursed with limitations of cognitive capacity or intelligence, and the conditions for learning from the game may vary in the course of the game. The mastery of a knowledge node will be indicated as a fraction or percentage, ranging from no mastery to full mastery. The players are also characterised by their motivation, which may change during the game. The following variables are required:

$I(i)$ The intelligence of player I (indicated as a fixed fraction)

$L(i,j)$ The learning effectiveness of player i during activity j (indicated as a fraction)

$P(i,j)$ The mastery level of player i in knowledge node j (indicated as fraction)

$P_c(i,j_c)$ Player i 's mastery level of knowledge in the child node j_c of parent node j

$M(i,j)$ The motivation of player i during activity j .

4.3.2. *Knowledge Deficits*

When the player moves to a new game activity j , the player is supposed to meet the activity's entry requirements (expected prior knowledge). Given the knowledge nodes covered by the activity, the prior knowledge requirement will comprise the sufficient mastery of all associated child nodes. Any deficits in the prior knowledge may affect the player's progress as the activity may be too demanding: the challenge will be too high, whereby motivation and performance would go down. Therefore, the player's knowledge mastery should always be compared with the prior knowledge requirements at hand, thus identifying any knowledge deficits upon entering an activity:

$D(i,j)$ The prior knowledge deficit of player i for addressing knowledge node j

When entering a new knowledge node the prior knowledge requirement would be the full mastery of the associated child nodes. The knowledge deficit $D(i,j)$ is thus expressed as a ratio through a summation over child nodes (for reasons of simplicity the influence of grandchild nodes is neglected).

$$D(i, j) = 1 - \frac{1}{N_c} \sum_{j_c=1}^{N_c} P_c(i, j_c) \quad (3)$$

4.4. Defining model dynamics

The overall relational model is depicted in figure 2.

[FIGURE 2 ABOUT HERE]

Figure 2. Causal model of learning upon engaging in game activity.

The player's state of knowledge mastery (1) as well as the complexity of the game activity (2) determine the severity of the challenge (3). According to flow theory, the severity of the challenge will influence the player's motivation (4), either by boredom or frustration. Also the attractiveness of the game activity (5) would influence motivation. Finally, both the player's motivation level and the player's intelligence (6) will determine the effectiveness of learning (7) and eventually the knowledge gained (8). These relationships translate into a set of equations.

First, the challenge associated with the activity is dependent on the knowledge state of the player. On the one hand a knowledge deficit would imply a positive challenge, on the other hand the player's knowledge level may exceed the requirements, for instance when a similar activity has been done before, thus posing a negative challenge. The challenge $C_H(i, j)$ regarding knowledge node i in activity j can be expressed and mapped onto the interval $[-1, 1]$ as follows:

$$C_H(i, j) = C(j) \cdot \left(e^{\frac{D(i)}{D(i)-1}} - e^{\frac{P(i)}{P(i)-1}} \right) \quad (4)$$

The severity of the challenge is assumed to be proportional with the complexity $C(j)$ of the activity. A deficit of prior knowledge $D(i)$ will contribute to a positive value of the challenge, but it can be counteracted by the (partial) mastery $P(i)$ of the knowledge nodes

addressed, which reduces the challenge by a negative contribution (the second term in equation 4). A perfect match of the task with the player's knowledge, namely no knowledge deficit and no mastery of the new knowledge node, would yield a challenge value equal to zero: the player is just ready to address the respective knowledge node.

As there are multiple knowledge nodes (N_k) addressed by an activity, the overall challenge is expressed as the arithmetic average to keep it on the interval [-1,1].

$$\bar{C}_H(j) = \frac{1}{N_k} \sum_{i=1}^{N_k} C_H(i, j) \quad (5)$$

Flow theory assumes optimal motivation in case of a perfect match. Imperfect matches will reduce motivation. This can be modelled with a normal distribution centred around the perfect match $C_H(i)=0$. The attractiveness $A(j)$ of activity j is added as a proportionality factor. Hence, the player's motivation in activity j is expressed as:

$$M(j) = A(j) \cdot e^{\left(\frac{-\bar{C}_H^2}{2\sigma_F^2}\right)} \quad (6)$$

Here σ_F is the flow factor, which is a scaling parameter indicating how sensitive the player's motivation is to a challenge mismatch. In case of a perfectly matching challenge ($C_H=0$), the player's motivation is maximal. For a non-zero challenge ($C_H \neq 0$, that is, a challenge either too easy or too hard) the flow factor σ_F determines how much the player's motivation is affected (in section 5.7, simulation study 6, the influence the flow factor will be further investigated). Herewith the motivation variable incorporates both the player's affective and cognitive states, which eventually determine the effectiveness of learning L . In fact, the learning effectiveness of player i in activity j is assumed to be proportional to both the player's motivation M and the player's intelligence I :

$$L(i, j) = M(i, j) \cdot I(i) \quad (7)$$

One adjustment should be allowed for including Vygotsky's principle of favouring challenges that are slightly beyond one's capabilities. If F_V is defined as the Vygotsky factor, denoting the required positive challenge value for optimised learning gains [44], equation 6 should then be rewritten as:

$$M(j) = A(j) \cdot e^{\left(\frac{-(\bar{C}_H - F_V)^2}{2\sigma_F^2}\right)} \quad (8)$$

4.5.Updating the player's knowledge states

After having successfully completed a game activity addressing one or more knowledge nodes the players' knowledge mastery states $P(i,j)$ should be updated. There are two elements. First, for each knowledge node j addressed by player i the knowledge mastery level will increase as new knowledge will be gained with an effectiveness of $L(i,j)$. The process of updating is described by the following recurrent expression:

$$P(i,j) = P(i,j) + L(i,j) \cdot (1 - P(i,j)) \quad (9)$$

Secondly, the process of mastering a parent node's knowledge is assumed to positively influence the knowledge mastery of child nodes, because the parent knowledge node is based on these and is supposed to integrate the child nodes' knowledge. For instance, doing a multiplication exercise would also increase one's fluency on numbers. Therefore, the process of mastering a parent node inherently contributes to the further mastery of all subordinate nodes in the full parent tree. This means that equation 9 also applies for updating the respective child node states. As a consequence, mastering a parent node in the game, be it partially, will indirectly reduce knowledge deficits of all conditional nodes in the knowledge hierarchy.

5. Simulation experiments

The serious gaming model described above was technically implemented using Scilab 5.5.2 (www.scilab.org). Table 1 summarises the independent and dependent variables of the computational model.

<<<INSERT TABLE 1 ABOUT HERE>>>

5.1. Baseline game and baseline player

For preliminary testing of the simulation a baseline game was generated based on a knowledge tree of 4 levels, presenting 3 child nodes for each parent. In accordance with equation 1, the resulting knowledge tree has 40 nodes. While allowing three occurrences of each knowledge node and preserving 3 knowledge nodes for each activity the 40 knowledge nodes support 40 game activities (equation 2). The knowledge nodes were randomly distributed over the game activities. In addition, each game activity was assigned a constant, moderate attractiveness of 0.5, as well as a constant, moderate complexity of 0.5. The activity with lowest overall rank in the knowledge tree was assigned the start activity and, conversely, the activity with highest overall rank in the knowledge tree was assigned the end activity. The process of playing the game would in principle allow any trajectory from the start activity across the remaining activities toward the end activity.

Likewise, a baseline game player was generated by assigning a moderate intelligence of 0.5 and lacking any prior knowledge with respect to the knowledge tree. The flow factor σ_F (equation 6) of this baseline player was set default to a value as large as 100 for initially disabling any flow effects. The Vygotsky factor was set to a default value of $F_V=0$. As a baseline strategy the player progresses through the game by selecting as the next activity the one that yields the highest learning effectiveness.

In the next sections 7 separate simulation studies will be presented, each focussing on a different variable.

5.2. Study 1, Fixed player: Stability across random game structure

The baseline game assumes a (random) distribution of knowledge nodes across the various game activities. Figure 3 shows the knowledge gain of the baseline player for 10 different (random) instances of the baseline game.

[FIGURE 3 ABOUT HERE]

Figure 3. Learning curves of the baseline player for 10 different (random) instances of the baseline game.

Here the player's knowledge gain (vertical scale) is expressed as the average mastery level of all nodes in the knowledge tree. It should be noted that, because each game instance in figure 3 has a different (random) distribution of knowledge nodes over game activities, these are all different games and should not necessarily produce the same learning curves or the same order of game activities. Obviously, the overall pattern of gradual knowledge gain is preserved. For each game activity, the spread of the knowledge values (vertical scale) can be expressed as a standard deviation. By averaging these standard deviations over all game activities (horizontal scale) a weighted value of the curves' vertical spread can be obtained, indicating the variability of learning curves. Figure 4 shows that the learning curve variability saturates to a level of 0.042 (standard deviation $SD=0.0002$), measured in the interval [500,1000]. Across these iterations the overall knowledge gained in the games was found to be stable at 0.938 ($SD=9*10^{-8}$).

[FIGURE 4 ABOUT HERE]

Figure 4. Learning curve variability as a function of the number of iterations.

Similar stable patterns were found for games with other numbers of knowledge nodes and activities.

5.3. Study 2, Fixed player: Attractiveness of game activities

In this study the attractiveness of game activities was varied. The baseline game was selected, while the uniform attractiveness of game activities was incrementally changed to cover the range [0,1]. The player was represented by the baseline player model. Figure 5 reveals that learning achievements increase with attractiveness of the activities. The simulation was then repeated with random knowledge distributions over the games' activities. The result was stable, showing standard deviations of the learning outcomes of typically 10^{-8} , too small to be even indicated in the figure.

[FIGURE 5 ABOUT HERE]

Figure 5. The knowledge gained versus the attractiveness of game activities.

Although attractiveness is included in the computational model as a simple multiplier (see equation 7), the shape of the curve in figure 5 is essentially non-linear. This can be attributed to the power law assumption of knowledge mastery that is enforced by equation 9: Re-iteration of equation 9 produces a (negative) power law that describes the asymptotic nature of the aim at perfect mastery.

Next, the attractiveness values were randomly drawn from a normal distribution (mean=0.5; SD=0.5), whereafter they were randomly assigned to the various game activities. The calculations show stable results from 150 iteration onward, in a pattern very similar to the one in figure 3. The learning curve variability was found to saturate to a level of 0.070 (SD=0.0003), measured in the interval [500,1000]. This saturation value is substantially higher than the value obtained with fixed attractiveness (figure 3). This can be attributed to the larger degree of freedom in the current case. The knowledge level attained after playing the game was found to be 0.936 (SD=0.025), which is very similar to the result found in study 1.

Hence, allowing for diversity of the knowledge nodes' attractiveness adds to the variability of learning during the game, but does not necessarily affect the overall knowledge gain.

5.4. Study 3, Fixed player: Influence of complexity of game activities

In this study the complexity of game activities was varied. The baseline game was selected, while the complexity of game activities was incrementally changed to cover the full interval [0,1]. A critical factor is the flow factor σ_F . Figure 6 shows the averaged knowledge gain results of 20 games for each of 100 complexity values (horizontal axis), for 6 different values of σ_F .

[FIGURE 6 ABOUT HERE]

Figure 6. The knowledge gained versus the complexity of game activities for 6 values of the flow factor σ_F .

The repetition of 20 games involved the random redistributing knowledge nodes over the games' activities. In accordance with empiricism, the figure reveals a downward tendency of the learning as complexity increases. Differences between the 6 curves are manifest. As can be understood from equation 6 a large value of σ_F neutralises the effects of challenge and thereby the direct influence of complexity: the curves with $\sigma_F \geq 0.5$ are quite fluent and the influence of complexity is weak. At smaller values of σ_F , however, the curves become more jagged. Here the player is more sensitive to mismatches between the player's knowledge level and the challenges offered. In accordance with flow theory the activity's complexity (see equation 4) comes into play by reducing motivation and the effectiveness of learning, which leads to a breakdown of the learning. The jagged patterns are persistent across larger numbers of iterations. The variability of the curves increases with smaller values of σ_F : standard deviations in figure 6 (vertical scale) increase from typically 1% at $\sigma_F=0.5$ up to 20% for $\sigma_F=0.01$. Moreover, the diversity of the curves increases when complexity is not assigned as a

fixed value for all knowledge nodes (as is the case in figure 6), but instead is distributed randomly across the nodes.

5.5.Study 4, Player intelligence

In this study the baseline game was used, while the intelligence of players was varied. Figure 7 shows how the overall knowledge gain depends on player intelligence. The horizontal axis (intelligence) is composed of 100 steps, each step being re-iterated with 20 baseline games with different distributions of knowledge nodes. The variability of the learning curves across games with different knowledge distributions is negligible, typically 10^{-7} .

[FIGURE 7 ABOUT HERE]

Figure 7. The knowledge gained versus player intelligence.

The curve in figure 7 (intelligence) is very similar to the one figure 5 (attractiveness). Both player intelligence and game attractiveness are included in the computational model in a similar way as simple multipliers (equations 6 and 7, respectively). In both cases the overall mastery is modulated by the power law assumption of knowledge mastery that is enforced by equation 9.

5.6.Study 5, Prior knowledge

In this study the baseline game was used, while the prior knowledge of the baseline player was varied. The prior knowledge gives the player a head start. In the calculations the prior knowledge was evenly distributed across the game's knowledge nodes. Figure 8 shows how the overall knowledge gain depends on the player's prior knowledge for each of the 100 horizontal coordinates, based on 20 iterations of games with different distributions of knowledge nodes.

[FIGURE 8 ABOUT HERE]

Figure 8. Player's knowledge as a function of prior knowledge, for different flow factors σ_F .

The flow factor σ_F was used as a mediating parameter, since prior knowledge has a similar role in determining the magnitude of the challenge as complexity has (see equation 4). For interpreting the curves of figure 8 one should take into account that there is little left for learning when the player's mastery is high from the start. Therefore one should consider the net knowledge gain, which is measured by the vertical distance to the dashed diagonal, as the latter reflect the knowledge level already mastered prior to playing the game. Although in particular the three upper curves suggest a rising tendency of the knowledge gained at increased prior knowledge, the nett effect is the opposite: the actual nett knowledge acquired decreases with prior knowledge. Drastic reduction of the learning occurs at smaller σ_F . When $\sigma_F=0.01$, which represents a player that is extremely sensitive to mismatches of challenges and knowledge level, the curve becomes highly irregular, showing some prior knowledge ranges of moderate learning alternated with ranges where the learning completely breaks down. The latter suggest catastrophic deadlocks in the game trajectories.

5.7.Study 6, Effects of flow

In this study the baseline game was played by the baseline player, be it that the flow factor changed to cover the whole interval $[0,1]$. Figure 9 shows the result for three complexity values (evenly distributed across game activities).

[FIGURE 9 ABOUT HERE]

Figure 9. The knowledge gained as a function of flow factor for three values of complexity.

For all complexity values the curves show that there is a critical point where the learning starts to break down. Critical points are in all cases well below $\sigma_F=0.1$ and they seem to move toward smaller σ_F at smaller complexity. Notwithstanding this breakdown, complexity weakly

modulates the learning outcomes across the whole range of flow factors σ_F . Similar curves were found for the dependences of prior knowledge, be it that the prior knowledge produces a knowledge off-set on the vertical scale. This similarity can be traced back to the role of both complexity and prior knowledge in determining the activities' challenges (equation 4).

5.8.Study 7, Playing strategies

In the experiments so far the players' strategy has been to progress through the game by selecting the activity with the highest learning effectiveness. Different decision strategies to traverse the game activities, however, might lead to different learning curves. Various options could be considered. For instance, the player might opt for selecting the activity that would produce the highest motivation as the next one. But as the player's intelligence is fixed, equation 8 shows that opting for the highest motivation would produce the same results as opting for highest learning effectiveness. Yet some other strategies so not coincide. The following strategies are compared:

A. Highest learning effectiveness

So far all studies have used this as the baseline strategy.

B. Balanced challenge

This strategy opts for a challenge closest to zero (see equations 4 and 5).

C. Minimal knowledge deficit

With this strategy the selection of the next game activity is based on minimal knowledge deficit.

D. Vygotsky's challenge

This strategy would select a positive challenge, which is slightly beyond one's capabilities, as indicated by the Vygotsky factor F_V .

E. Random choice

Here the player just randomly picks the next activity for the set of options.

F. Random strategy

Rather than adopting a fixed strategy the player randomly selects one of the top 4 strategies for each decision.

Importantly, the model only requires a strategy to be executed, but it remains indifferent about who or what is the strategy agent: the player, a teacher, the game or any other agent. For allowing these strategies to produce different outcomes, the baseline player was adapted to have a flow factor of $\sigma_F = 0.05$ rather than $\sigma_F = 100$. Also the heterogeneity of the baseline game was removed by adding random values of complexity and attractiveness to the game activities. These values were drawn from a normal distribution (mean=0.5 and SD=0.1). For the Vygotsky strategy the Vygotsky factor was changed from $F_V = 0$ to $F_V = 0.1$. Figure 10 shows the learning curves for the 6 strategies.

[FIGURE 10 ABOUT HERE]

Figure 10. learning curves for the 6 strategies A, B, C, D, E, F, respectively.

The dashed-dotted curves show the average result over 1000 iterations for the random choice strategy (E) and the random strategy (F). Standard deviations (vertical scale) for both sampled curves remained well below 0.1. First of all, this case reveals substantial differences between different strategies. Not just the learning curves are different, but also their endpoints, which reflect the knowledge state after game completion. Second, strategies A and D show a prosperous start, but the learning seems to saturate. In contrast, strategies B and C show smaller learning rates, but in the end produce better results. The winning strategy in this example game would be Vygotsky's strategy. However, the data only reflect one particular

case. Random choice is the weakest strategy, although in this case still quite some knowledge is gained. More investigations are needed to reproduce and establish the effects.

As a final step Table 2 presents the cross-tabulation of Kendall-tau distances between the 4 single run strategies (A, B, C, D).

<<<INSERT TABLE 2 ABOUT HERE>>>

The Kendall-tau rank distance uses the concordances between pairs of elements for comparing two sequences. Equal sequences would receive a distance of 0, opposite sequences receive a distance of 1, while a value of 0.5 indicates complete unrelatedness. From the table it can be seen that most distance values are close to 0.4, which indicates that the order of activities is very different across the different strategies. Hence, this study shows that the playing strategy or thereby the learning strategy are important factors that influence the order of activities as well as the final learning achievements.

6. Discussion and conclusion

The simulation model presented in this paper was motivated by the opportunities that it would offer to study detailed in-game mechanisms and processes related to the productive mastery of new knowledge and skills. With respect to the research questions framed in the introduction it is concluded that it proved possible to formulate an expressive computational model for the process of learning from games that avoids inherent complexity by taking up a meso-level perspective. The model turns out to deliver stable and reproducible outcomes that are consistent with empiricism in serious gaming and allows for investigating detailed and quantitative dependences between relevant concepts and variables. It was demonstrated how both attractiveness of the game and the player's intelligence support the knowledge gained

1 through a monotonously rising curve with a gradually diminishing effect. Also, the downward
2 tendency of the learning outcomes as complexity increases was observed. The player's
3 sensitiveness for mismatches between the player's capabilities and the challenges offered was
4 identified as a crucial factor for productive learning, which in agreement with empiricism and
5 directly related to flow theory. Finally, it was observed how different strategies may lead to
6 different learning curves and eventually different outcomes.

7
8
9
10
11
12
13
14 The model as much as the methodology are presented and intended as a starting point
15 for establishing a new line of computational research on serious gaming. Its potential is in
16 obtaining a deeper understanding of the interplay of learning and playing in serious games.
17 Eventually, this research could lead to simulation-based methods and tools that directly
18 support serious game designers at optimising their games' effectiveness for learning. As the
19 model was deliberately kept simple, it offers various opportunities for extension. First,
20 although the model is based on concepts and theories of learning, it so far omits instructional
21 content and didactics. Instead it postulates that the engagement in a game activity constitutes
22 a learning experience that has some effectiveness at gaining new knowledge or skills. On the
23 expense of increased complexity, various instructional concepts, e.g. guidance, feedback,
24 testing or reflection, could be included in the model to enhance model fidelity and precision.
25
26 Second, the model does not include cognitive models of human information processing and
27 human learning, but instead just relies on the phenomenology of the process of play. Linking
28 the model with existing architectures of human cognition, e.g. Soar [49], Clarion [50], ACT-R
29 [51] or COGENT [52], would allow for including psychological constructs such as fatigue,
30 attention, intention, concentration, meta-cognition, emotion, perception, cognitive load,
31 processing capacity, response times and some more. Third, in the model presented in this
32 paper the human player is only characterised by prior knowledge and intelligence, while a lot
33 of additional concepts and indicators might be used to describe and qualify the player. Today,

1 advances in datamining, learning analytics and a wide variety of sensor technologies greatly
2 extend the opportunities for collecting and extracting user profile data. Fourth, the key
3 strategies investigated in this study have their limitations. Particularly, they are used as fixed
4 strategies, whereas real players might change their strategies during the game because of
5 changed conditions or changed moods. Also, by relying on the Traveling Salesperson
6 Problem as a metaphor of progressing thorough a game, the players were restrained from
7 revisiting activities. Still, revisiting an activity node after initial failure could be a sensible
8 strategy. In all cases the strategies used reflect opportunism by favouring decisions that offer
9 the largest utility on the short term: no thinking ahead is occurring. Such strategy of local
10 optimisation may eventually not produce the best overall learning outcome. Fifth, as the
11 model presupposes well-defined knowledge structures (e.g. skills hierarchies, competence
12 maps, goal structures), its extension to ill-defined domains such as soft-skills poses a
13 challenge. Still, in those cases machine learning approaches may be applied for identifying
14 knowledge patterns and their mapping on to game activities and behaviours [46].

15
16 Finally, the findings of this study are well in agreement with intuitions and
17 empiricism, but extending this research and linking the model with real games would allow
18 for more strictness and empiricial validation, as well as enhanced theory development, tuned
19 model extensions and deeper insights in game-based learning. Eventually, computational
20 modelling could assist game design by testing structural pathways and identifying
21 weaknesses, altogether leading to serious games that are more effective for learning.

22 **7. Acknowledgement**

23 This research was partially funded by the European Union's Horizon 2020 research
24 and innovation programme under grant agreement No 644187, the RAGE project
25 (www.rageproject.eu).

8. References

- [1] D. Michael, S. Chen, Serious games: games that educate, train and inform, Thomson Course Technology, Boston, 2006.
- [2] E. Klopfler, S. Osterweil, K. Salen, Moving learning games forward; obstacles, opportunities & openness, MIT-The Education Arcade, Boston, 2009. Retrieved from http://education.mit.edu/papers/MovingLearningGamesForward_EdArcade.pdf.
- [3] C. Aldrich, The Complete Guide to Simulations and Serious Games: How the Most Valuable Content Will be Created in the Age Beyond Gutenberg to Google, Pfeiffer, San Francisco, 2009.
- [4] M.D. Dickey, Engaging by design: How engagement strategies in popular computer and video games can inform instructional design, Educational Technology Research and Development 53-2 (2005) 67-83.
- [5] J.P. Gee, What videogames have to teach us about learning and literacy, Palgrave Macmillan, New York, 2003.
- [6] M. Polanyi, The Tacit Dimension, University of Chicago Press, Chicago, 1996.
- [7] S. Björk, J. Holopainen, Patterns in Game Design, Charles River Media, Hingham, 2005.
- [8] A. De Gloria, F. Bellotti, R. Berta, Serious games for education and training. International Journal of Serious Games, 1-1 (2014). Retrieved from <http://journal.seriousgamessociety.org/>.
- [9] W. Westera, R. Nadolski, H. Hummel, I. Wopereis, Serious Games for Higher Education: a Framework for Reducing Design Complexity, Journal of Computer-Assisted Learning, 24-5 (2008) 420-432.
- [10] K. Salen, E. Zimmerman, Rules of Play, Game Design Fundamentals, MIT Press, Cambridge MA, 2004.
- [11] W. Westera, Games are motivating, aren't they? Disputing the arguments for digital game-based learning, International Journal of Serious Games, 2-2 (2015). Retrieved from <http://journal.seriousgamessociety.org/>.
- [12] S. De Marchi, Computational and Mathematical Modelling in the Social Sciences, Cambridge University Press, New York, 2005.
- [13] C. Achen, Empirical implications of theoretical modelling (Report of the Political Sciences Programme, Directorate for Social, Behavioral and Economic Sciences), National Science Foundation, Arlington, 2005.

- [14] A. Mezzacappa, Top Breakthroughs in Computational Science [Special Issue], Scientific Discovery through Advanced Computing 11 (2009) 32-33. Retrieved from <http://www.scidacreview.org/0901/pdf/bt.pdf>.
- [15] T. Reeves, Design research from a technology perspective, in: J. Van den Akker, K. Gravemeijer, S. McKenney, N. Nieveen (Eds.), Educational Design Research, Routledge, New York, 2006, pp. 52-66.
- [16] W. Van der Vegt, W. Westera, E. Nyamsuren, A. Georgiev, I Martínez Ortiz, RAGE Architecture for Reusable Serious Gaming Technology Components. International Journal of Computer Games Technology, 2016, Advance online publication, doi:10.1155/2016/5680526. Retrieved from <http://www.hindawi.com/journals/ijcgt/2016/5680526/>.
- [17] A. Bates, Technology, Open Learning and Distance Education, Routledge, London/New York, 1991.
- [18] A. Bates, Managing Technological Change: Strategies For College And University Leaders, Jossey Bass/John Wiley, San Francisco, 2000.
- [19] CERI, Knowledge Management in the Learning Society, OECD - Organisation for Economic Co-operation and Development, Paris, 2000.
- [20] Westera, W. (2012). The eventful genesis of educational media. Education and Information Technologies, 17(3), 345-360.
- [21] A. Gelman, P-values and statistical practice, Epidemiology 24-1 (2013) 69-72.
- [22] V.E. Johnson, Revised standards for statistical evidence. PNAS - Proceedings of the National Academy of Sciences 110-48 (2013) 19313–19317. doi: 10.1073/pnas.1313476110. Retrieved from <http://www.pnas.org/content/110/48/19313>
- [23] T.M. Connolly, E.A. Boyle, E. MacArthur, T. Hailey, J.M. Boyle, A systematic literature review of empirical evidence on computer games and serious games, Computers & Education 59-2 (2013) 661–686. doi: 10.1016/j.compedu.2012.03.004.
- [24] C. Linehan, B. Kirman, S. Lawson, G. Chan, Practical, appropriate, empirically-validated guidelines for designing educational games, in: CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, 2011, pp. 979-988. doi: 10.1145/1978942.1979229.
- [25] S. Schaeffer, Tic-Tac-Toe (Naughts and Crosses, Cheese and Crackers, etc.), Mathematical Recreations website, 2002. Retrieved from <http://www.mathrec.org/old/2002jan/solutions.html>.

- [26] U. Neisser, G. Boodoo, T.J. Bouchard Jr, A.W. Wade Boykin, N. Brody, N., S.J. Ceci, D.F. Halpern, J.C. Loehlin, R. Perlo, R.J. Sternberg, S. Urbina, Intelligence: knowns and unknowns, *American Psychologist* 51-2 (1996) 77–101.
- [27] R.J. Sternberg, J.C. Kaufman, Human abilities, *Annual Review of Psychology* 49 (1998) 479–502.
- [28] P.R. Pintrich, D.H. Schunk, *Motivation in education: theory, research and applications*, Prentice Hall, Englewood Cliffs NJ, 1996.
- [29] R.M. Ryan, E.L. Deci, Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being, *American Psychologist* 55-1 (2000) 68-78.
- [30] J.M. Keller, Development and Use of the ARCS Model of Motivational Design, *Journal of Instructional Development* 10-3 (1987) 2-10.
- [31] J.M. Keller, First principles of motivation to learn and e3-learning, *Distance Education* 29 (2008) 175-185.
- [32] L.W. Anderson, D.R. Krathwohl, *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, Longman, New York, 2001.
- [33] B.S. Bloom, M.D. Engelhart, E.J. Furst, H.W. Hill, D.R. Krathwohl, *Taxonomy of educational objectives: The cognitive domain*, Longman, New York, 1956.
- [34] P.R. Pintrich, D.R. Cross, R.B. Kozma, W.J. McKeachie, *Instructional psychology*, *Annual Review of Psychology* 37 (1986) 611–651.
- [35] R.C. Schank, T.R. Berman, K.A. Macpherson, Learning by doing, in: C.M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory*, vol 2, Lawrence Erlbaum Associates, Mahwah NJ, 1999, pp. 161-181.
- [36] G. Bickley, Personality traits, learning strategies, and performance, *European Journal of Personality* 10 (1996) 337–352.
- [37] P. Ekman, W.V. Friesen, *Facial action coding system: Investigator's guide*, A Human Face, Douglas AZ, 1978.
- [38] C. Bateman, R. Lowenhaupt, L.E. Nacke, Player typology in theory and practice, in: *Proceedings of the 2011 DiGRA International Conference : Think Design Play*, Digital Games Research Association, Utrecht, 2011.
- [39] R.M. Felder, L.K. Silverman, Learning and Teaching Styles in Engineering Education, *Engineering Education* 78-7 (1998) 674–681. Retrieved from <http://www.ncsu.edu/felderpublic/Papers/LS-1988.pdf>.

- [40] R.M. Felder, Are learning styles invalid? (Hint: No!), On-Course Newsletter Sept. 27 (2010). Retrieved from [http://www4.ncsu.edu/unity/lockers/users/f/felder/public/Papers/LS_Validity\(On-Course\).pdf](http://www4.ncsu.edu/unity/lockers/users/f/felder/public/Papers/LS_Validity(On-Course).pdf)
- [41] D.A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, Prentice-Hall, Englewood Cliffs NJ, 1984.
- [42] S. Buckingham Shum, R. Deakin Crick, Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics, in: *Proceedings of the 2nd International Conference on Learning Analytics & Knowledge*, ACM, New York, 2012, pp. 92-101. doi: 10.1145/2330601.2330629.
- [43] M. Csikszentmihalyi, *Flow: The psychology of optimal experience*, Harper Perennial, New York, 1991.
- [44] L.S. Vygotsky, *Mind in society. The development of higher psychological processes*, Harvard University Press, Cambridge MA, 1933/1987.
- [45] J. Heller, C. Steiner, C. Hockemeyer, D. Albert, Competence-based knowledge structures for personalised learning, *International Journal on E-learning* 5-1 (2006) 75-88.
- [46] V.J. Shute, M. Ventura, *Measuring and supporting learning in games: Stealth assessment*, The MIT Press, Cambridge MA, 2013.
- [47] M.M. Flood, The Traveling-Salesman Problem, *Operations Research* 4 (1956) 61-75.
- [48] F.J.R.C. Dochy, *Assessment of Prior Knowledge as a Determinant for Future Learning: The use of prior knowledge state tests and knowledge profiles*, Lemma BV, Utrecht/London, 1992.
- [49] J.E. Laird, *The Soar Cognitive Architecture*, MIT Press, Cambridge, MA, 2012.
- [50] R. Sun, The CLARION cognitive architecture: Extending cognitive modeling to social simulation, in: R. Sun (Ed.) *Cognition and Multi-Agent Interaction*, Cambridge University Press, New York, 2006, pp. 79-102.
- [51] J.R. Anderson, C. Lebière, *The atomic components of thought*, Erlbaum, Mahwah NJ, 1998.
- [52] R. Cooper, P. Yule, J. Fox, D. Sutton, COGENT: An environment for the development of cognitive models, in: U. Schmid, J. F. Krems & F. Wysotzki (Eds.), *A Cognitive Science Approach to Reasoning, Learning and Discovery*, Pabst Science Publishers, Lengerich (Germany), 1998, pp. 55-82.

Highlights

- Cognitive flow in serious games can be effectively modelled in computational models
- Complexity of learning from games can be understood from activity-oriented computational models
- Models explain how different user strategies in serious games produce different learning curves
- Computational models open up a new research field of learning processes in serious games

Changes

All issues solved.

REVIEWER 1

1. The abstract should be punchier, really in a nutshell outlining the major achievements of the paper, although the first two are there, in my view the authors should replace the last sentence outlining the importance of the proposed approach.

SOLVED

The final sentence of the abstract was reformulated in order to stress the importance of the work: "The model allows researchers to deeply investigate quantitative dependences between relevant game variables, gain deeper understanding of how people learn from games, and develop approaches to improving serious game design."

REVIEWER 1

2. All sections and subsections should be numbered. It is then easier to refer to the sections.

Reviewer 1

SOLVED

All sections numbered in accordance with journal style.

REVIEWER 1

3. In Study 7 section, it should be:

"Figure 10, learning curves for the 6 strategies A, B, C, D, E, F respectively. (Just take the brackets off).

SOLVED

REVIEWER 2

1. The usage of the term "gameplay" throughout the paper is a bit ambiguous. Within the computer game community, the term gameplay is used to describe the specific way players interact with a game in the sense of game mechanics, i.e. the formal rules of that game and user interface. Throughout the paper the term is used more in the sense of "the process of playing a game". Additionally, the term is not consistently written as "gameplay", but sometimes as "game play". Therefore, please avoid using this term and check the paper for consistent usage. This already affects the title and the abstract of the paper. Please rename the title towards "Simulating how People Learn while playing Serious Games: A Computational Modelling Approach" or similar.

SOLVED

Replaced "game()play" with "playing the game"
Title adjusted accordingly

REVIEWER 2

2. The chapter "Model Elaboration", sub-chapter "Defining the Knowledge Model", is difficult to understand due to mixing-up terminology. Using graph theoretical terminology, what you are describing is a rooted tree, not a hierarchical network. Particularly, it is a perfect k-ary tree, where each node has k children, except the leaf nodes, which have zero children, and all leaf nodes are at the same depth. Please, avoid the term "generation" when talking about such trees, instead use "level" or "depth". Also, there is no "symmetrical" hierarchy. Additionally, without giving the precise definition of what kind of tree is used the underlying simplification (i.e. knowledge is represented by a perfect k-ary tree) cannot be understood. In your terms $N_c = k$ and number of generations = depth). It would help to give references to understand that model. Also, there is a strong simplification of using perfect k-ary trees instead of any rooted trees which should be motivated.

SOLVED

The description was adjusted in order to consistently comply with graph terminology (rooted tree, k-ary tree, levels). An annotation was included explaining that in educational sciences and cognitive psychology other terms are commonly being used (e.g. Blooms learning objectives/skills taxonomies are labelled as hierarchical models in cognitive, affective and sensor-motor domains).

Additional text was included to explain why perfect k-ary trees are used (section 4.1):

"For reasons of simplification the knowledge model is described as a perfect k-ary tree, where k is called the degree, which is the number of child nodes of each parent node. Assuming a perfect k-ary tree, however, does not imply a fundamental constraint, as it would still allow for transforming it into any irregular tree model by removing selected nodes and edges from the tree. "

REVIEWER 2

3. The following chapter "Defining the Game Model" is also a bit awkwardly explained. In my understanding, game activities are elements of a finite set of game activities. There are no relations between the game activities such as precedence. The set of nodes of the knowledge tree is mapped on the set of game activities in such a way that the number of nodes mapped on a single game activity is equal for all game activities (described by N_i), and each node supports the same number of game activities (described by N_k). Otherwise, equation (2) will not hold. Now by selecting N_n , N_i , and N_k the number of activities N_a is calculated, and N_i and N_k have to be selected carefully so that N_a becomes a natural number. Again, this is a simplification which should be motivated.

SOLVED

-Additional text was included in section 4.2

"Although, in principle, the mapping relationships between game activities and knowledge nodes will be diverse, we assume for reasons of simplicity that each game activity covers a fixed number of knowledge nodes, and that the number of occurrences of the knowledge nodes is likewise a fixed number."

and

"This simplification does not pose a basic restriction of the proposed computational model as it only affects the model's input data."

-Precedence relationships are referred to in section 3.5 by the prior knowledge requirements and the game narrative, which may impose a restricted, logical or causal order of game events, and thereby would even reduce the complexity of the model.

REVIEWER 2

4. With the identified simplifications the whole model is very special and probably not comparable to models reflecting real Serious Games. When using this approach to model the playing of a specific Serious Game one wouldn't expect a perfect k-ary tree as the knowledge model and the number of nodes per activity and vice versa is not expected to be the number for each node and each activity. This must be made explicit in the discussion section

SOLVED

This is covered by the previous 2 actions.

REVIEWER 2

5. On page 10 it is stated: "Achieving cognitive flow requires a cautious adjustment of the challenges offered to the player's abilities." Therefore, cognitive flow is a mental state of the player. Within the model, the player's action (performance?) should be based on that mental state. Surprisingly, the "flow factor" in the model is the standard deviation of the distribution of the player's motivation to a mismatch. It is not obvious how the concept of flow (mental state of a person), which is dynamic - depending on the challenge of an activity and the current ability of the player - can be static input data. Obviously, the motivation (equation 6) has something to do with the flow, but for my understanding of the model objectives, flow must also be calculated/updated after the completion of each game activity.

SOLVED

The text below equation(6) was extended to clarify the role of the flow factor σ_F :

"Here σ_F is the flow factor, which is a scaling parameter indicating how sensitive the player's motivation is to a challenge mismatch. In case of a perfectly matching challenge ($C_H=0$), the player's motivation is maximal. For a non-zero challenge ($C_H \neq 0$, that is, a challenge either too easy or too hard) the flow factor σ_F determines how much the player's motivation is affected (in section 5.7, simulation study 6, the influence the flow factor will be further investigated)."

REVIEWER 2

6. Grammar and syntax

SOLVED

- cf." removed and/or replaced with "see"
- "viz." replaced with "namely" or "that is"
- All personal pronouns ("we") removed
- "Game theory" is correct: rational decisions under conditions of uncertainty; references are in place. Explanation is in section 2.1.
- "Pave the way for", replaced with "would allow for"
- Typo "expertise that are required" Replaced with "expertise that is required."
- Typo: "verified for as explaining" replaced with "verified for explaining"
- Typo: assed the thousands separator used in numbers
- "... should stay away from ..." by "... should avoid ...".
- "of a larger grainsize" removed
- "etcetera)" removed
- "large grainsize" replaxced by "high level".
- "Thus, each game activity offers ..." instead of "Each game activity thus offers ...".
- "... some of the expected learning ..." instead of "... some of expected learning ...".
- "from" removed from "suffering from substantial ..."
- "complete mapping" replaced with "unambiguous"
- "Therefore" replaced with "Therefore,"
- "positive dependence" reformulated: "Figure 5 reveals that learning achievements increase with attractiveness of the activities."
- "[0,1]" instead of "[0.1]"
- "..., be it that still quite ..." replaced with "..., although in this case still quite..."
- Typo: Reference 24, page numbers corrected.

REVIEWER 3

1. How could the computational model add value and support a good design if one or many of the elements in the model are omitted? Are there any recommendations to the designers on how they can use the computational model to get the best results?

SOLVED

The main purpose of the model is "...to allow and support researching what happens during playing a game under a variety of conditions" (see section 1.3)

This was further emphasised In section 6 by the following extended text:

"The model as much as the methodology are presented and intended as a starting point for establishing a new line of computational research on serious gaming. Its potential is in obtaining a deeper understanding of the interplay of learning and playing in serious games. Eventually, this research could lead to simulation-based methods and tools that directly support serious game designers at optimising their games' effectiveness for learning. "

REVIEWER 3

2. Assuming the computational model is for researchers in the area of games and learning, are there other target audiences and how would the model benefit them - assuming a no. of designers are not perhaps interested in all the elements in the model? Identifying the contributions of the work to a broader audience would strengthen the paper.

SOLVED

See previous items.

REVIEWER 3

3. The paper is very conceptual and does not provide any examples. How would the model perform in different learning domains? For example, a domain or types of knowledge that are very clear (e.g. procedural knowledge) and one that has more complex or soft competences would be harder to model (e.g. project management). Some discussions about the usage of the model and it's suitability for learning in different domains or contexts would also strengthen the paper and the model would appear more convincing.

SOLVED

In section 6 the following text was added:

“Fifth, as the model presupposes well-defined knowledge structures (e.g. skills hierarchies, competence maps, goal structures), its extension to ill-defined domains such as soft-skills poses a challenge. Still, in those cases machine learning approaches may be applied for identifying knowledge patterns and their mapping on to game activities and behaviours [46].”

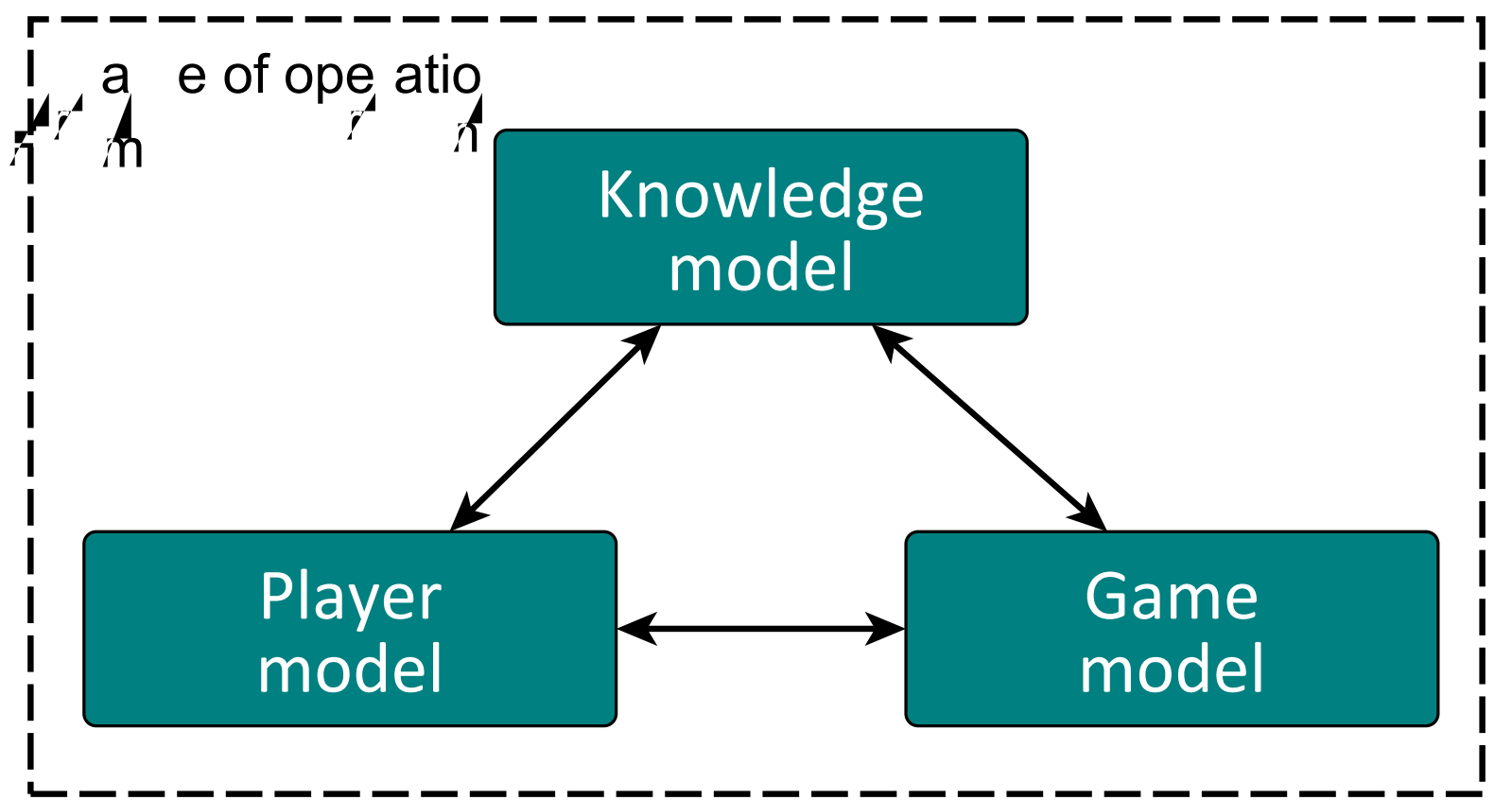
	Independent variables	Dependent variables
Knowledge tree	<ul style="list-style-type: none">• Knowledge generations• Child degree	
Game	<ul style="list-style-type: none">• Distribution of knowledge nodes over activities• Attractiveness of each activity• Complexity of each activity	
Player	<ul style="list-style-type: none">• Intelligence• Prior knowledge• Playing strategy• Flow factor	<ul style="list-style-type: none">• Acquired knowledge for each activity• Challenge of each activity• Motivation for each activity• Learning effectiveness for each activity• Game trajectory

Table 1. Overview of variables in the computational model

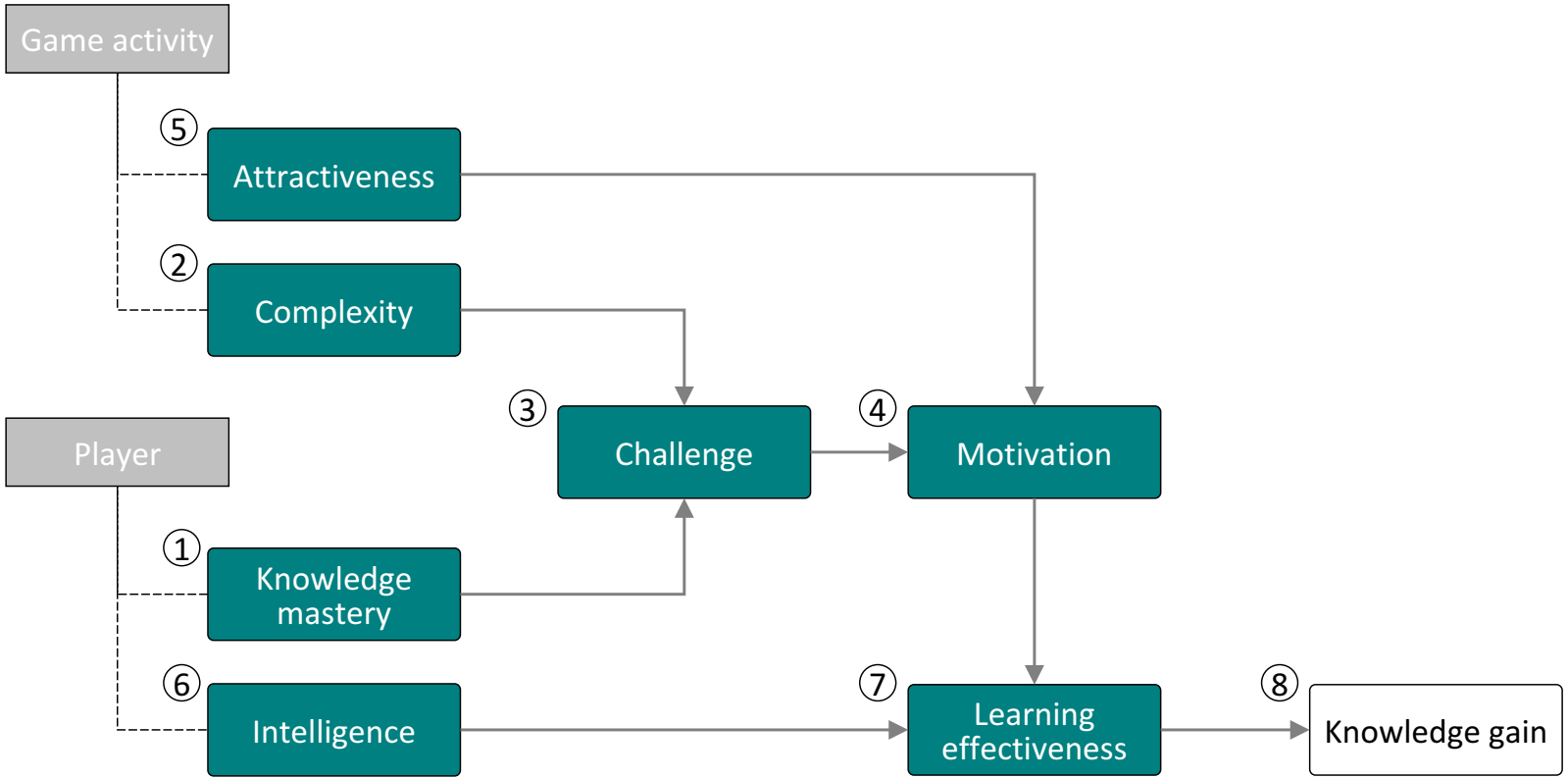
	A	B	C	D
A. Learning effectiveness	0.0	0.341	0.435	0.347
B. Balanced challenge	0.341	0.0	0.181	0.373
C. Knowledge deficit	0.435	0.181	0.0	0.423
D. Vygotsky	0.347	0.373	0.423	0.0

Table 2. Kentall-tau rank distances between the activity orders of 4 strategies.

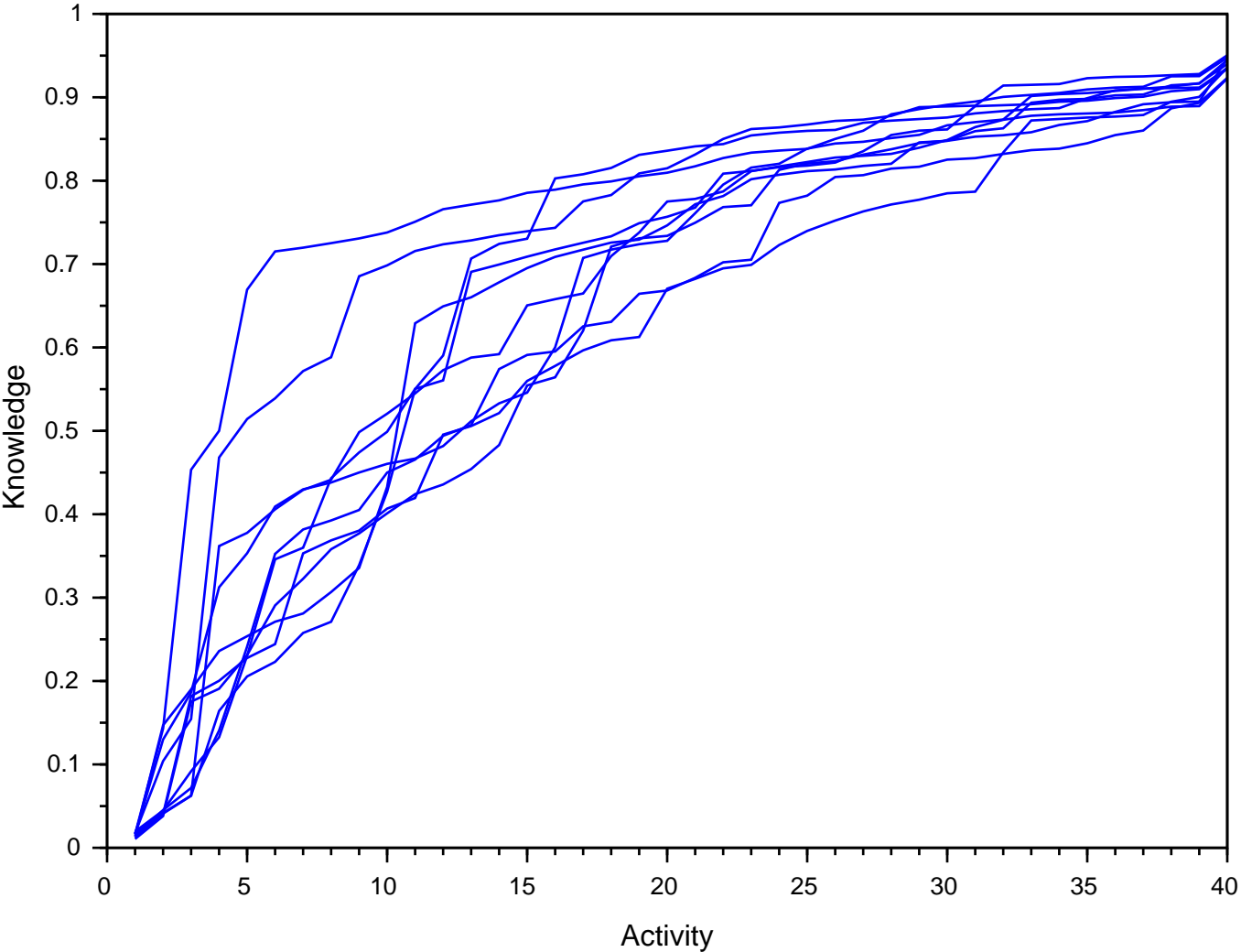
Figure



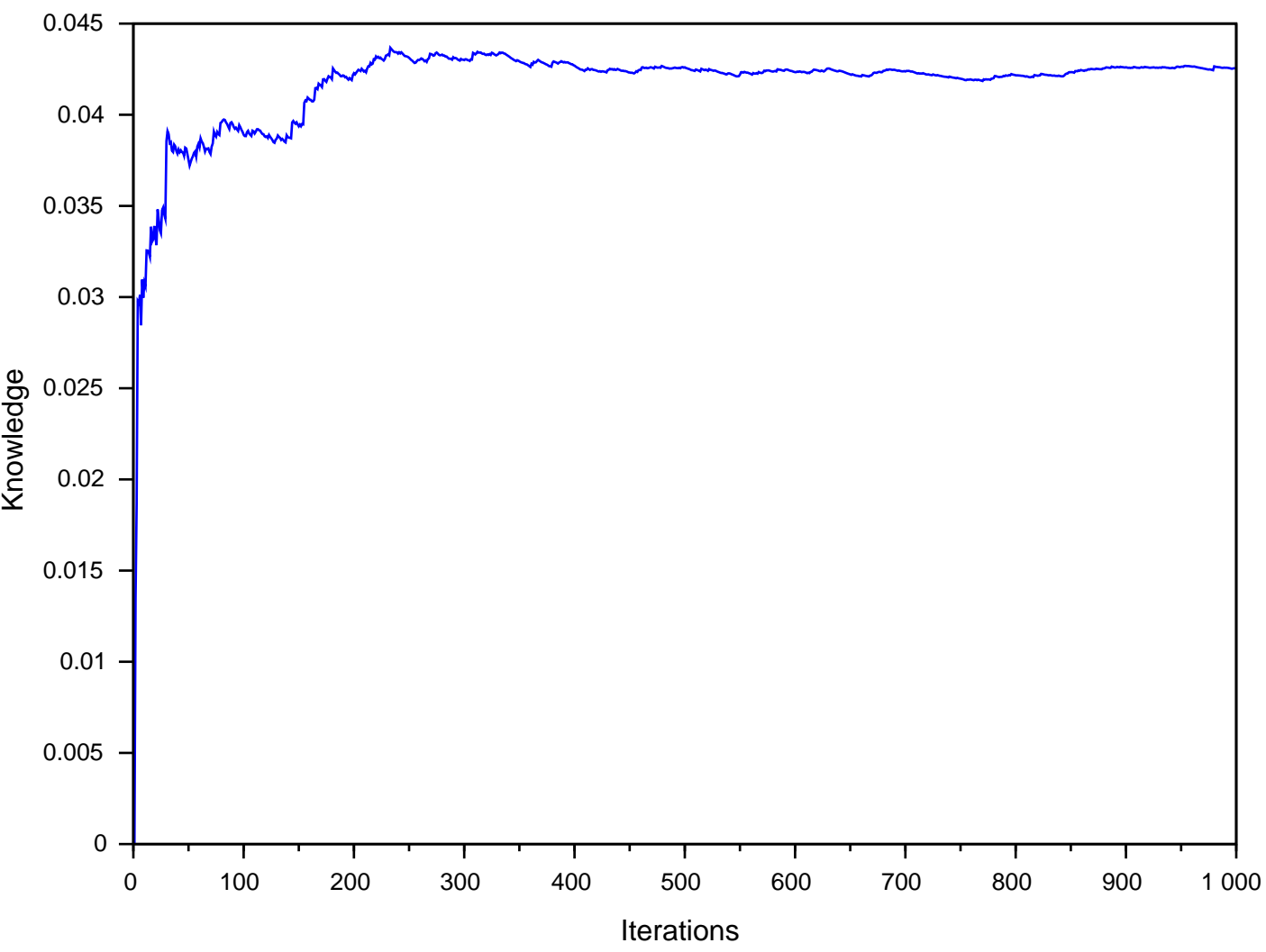
Figure



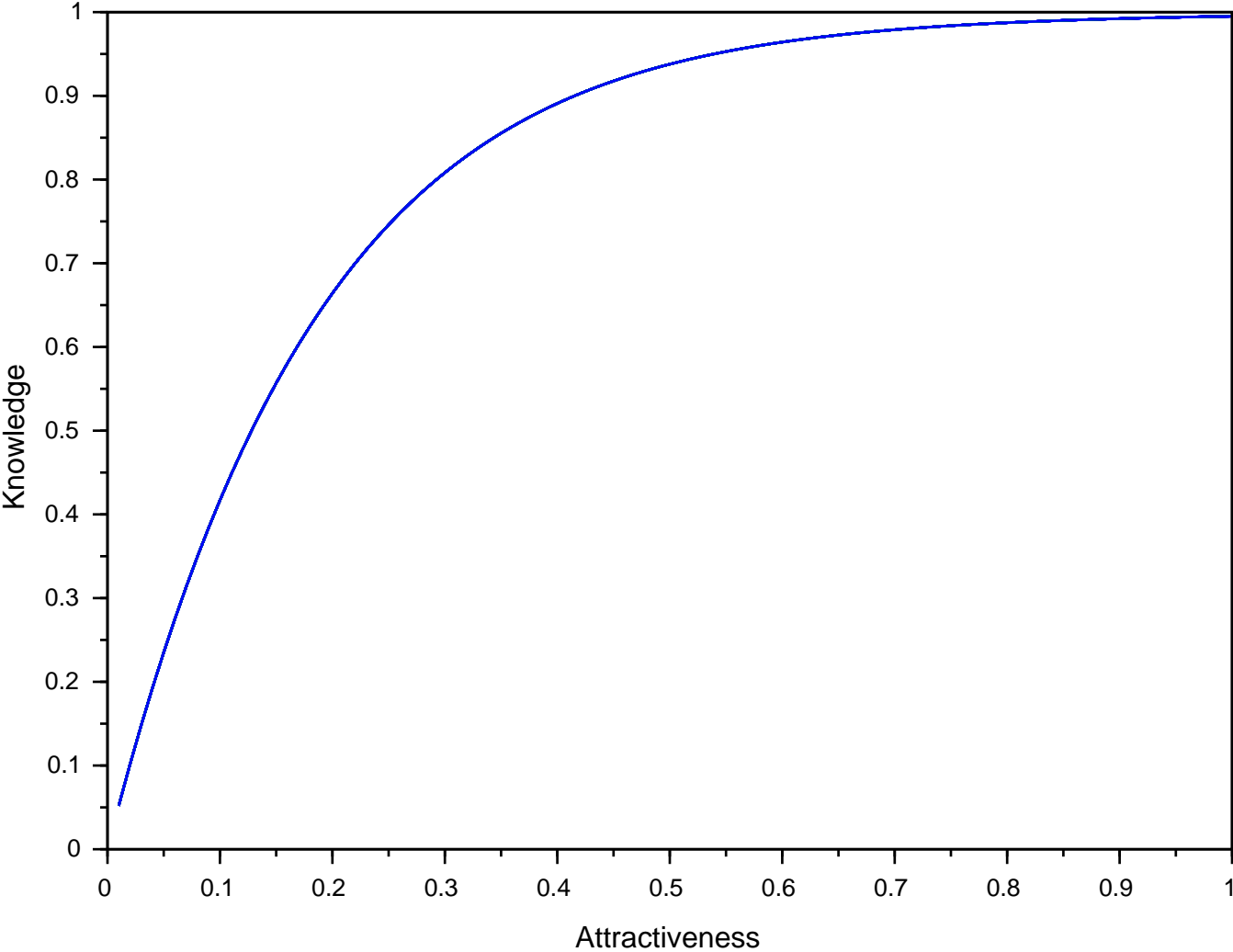
Figure



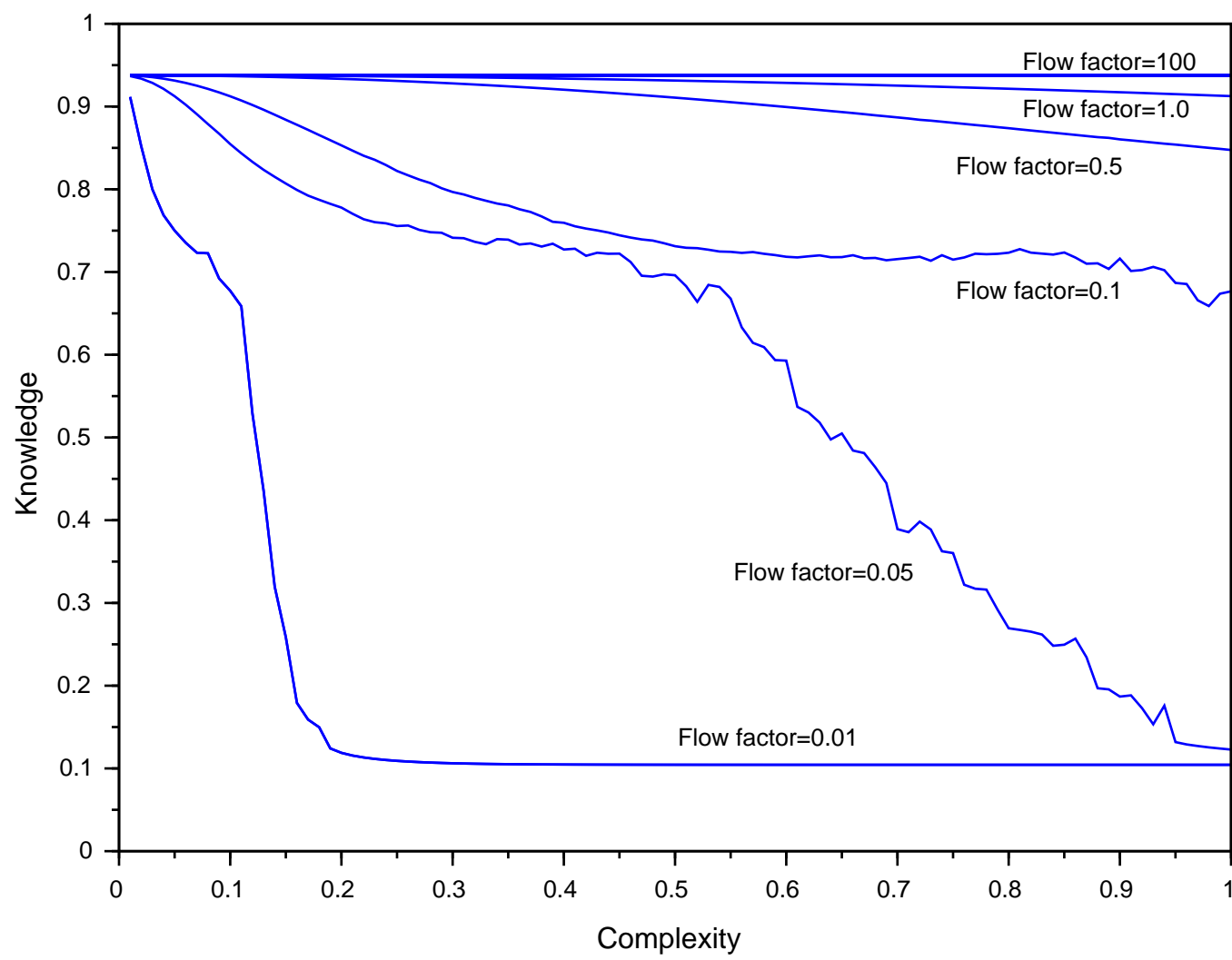
Figure



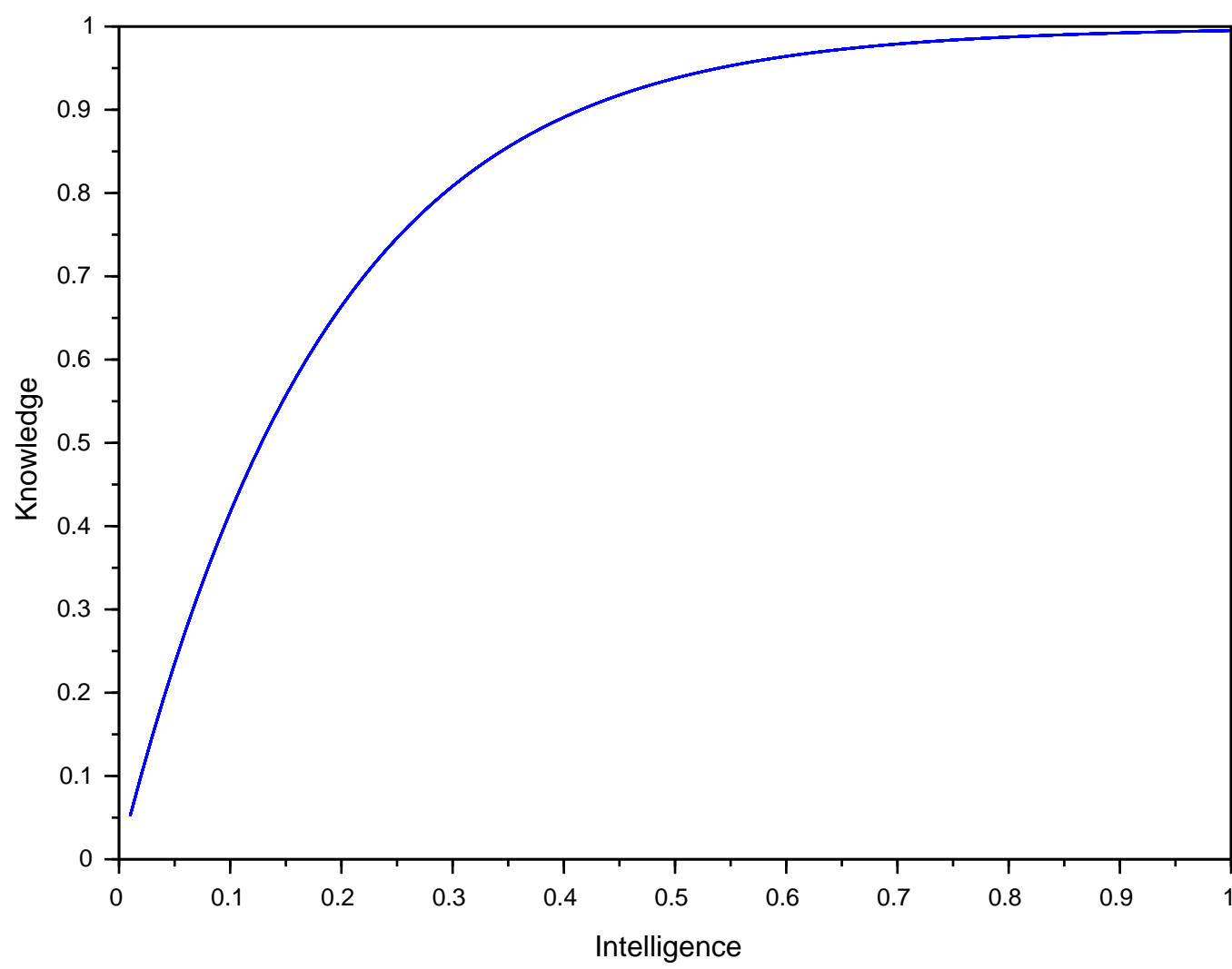
Figure



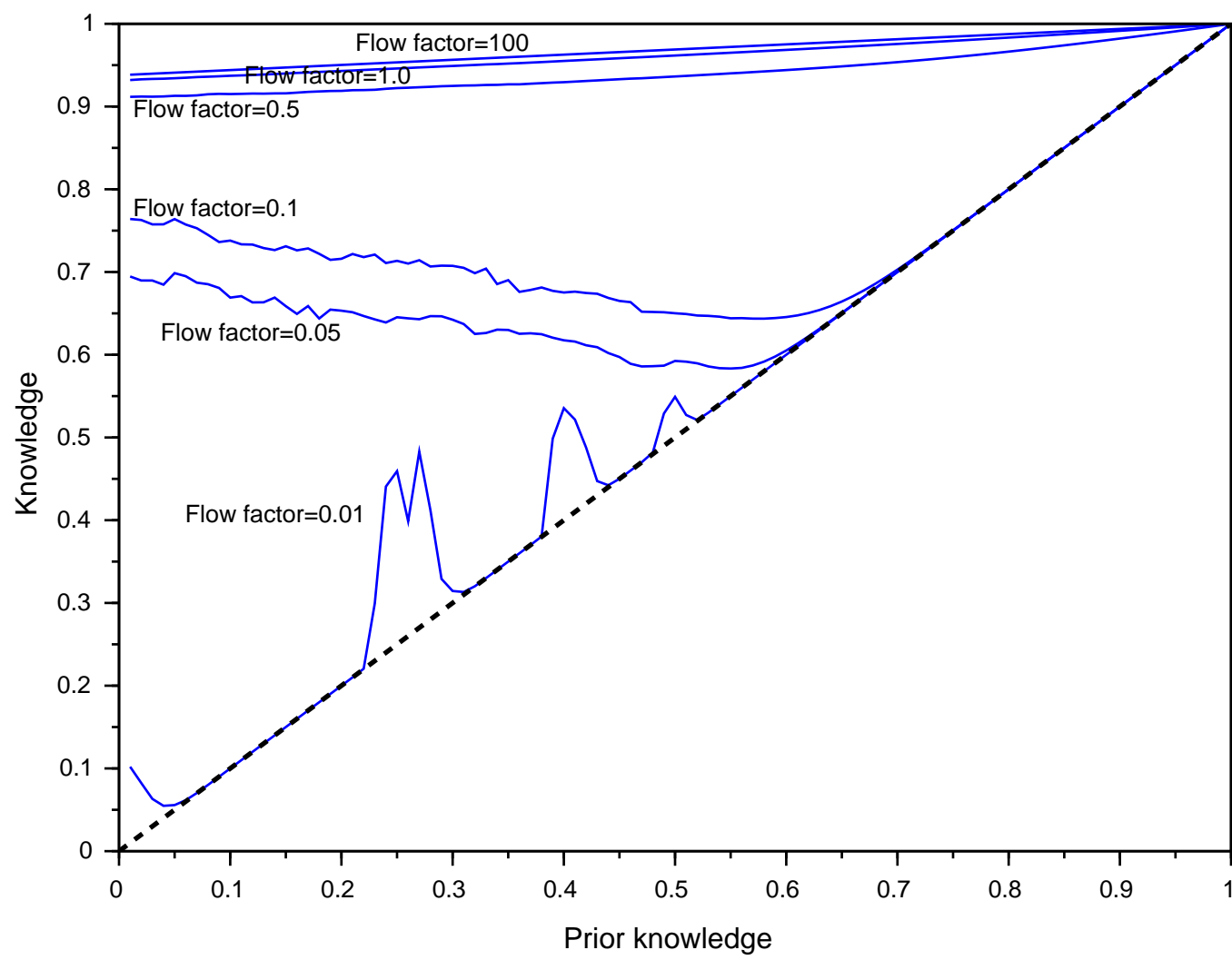
Figure



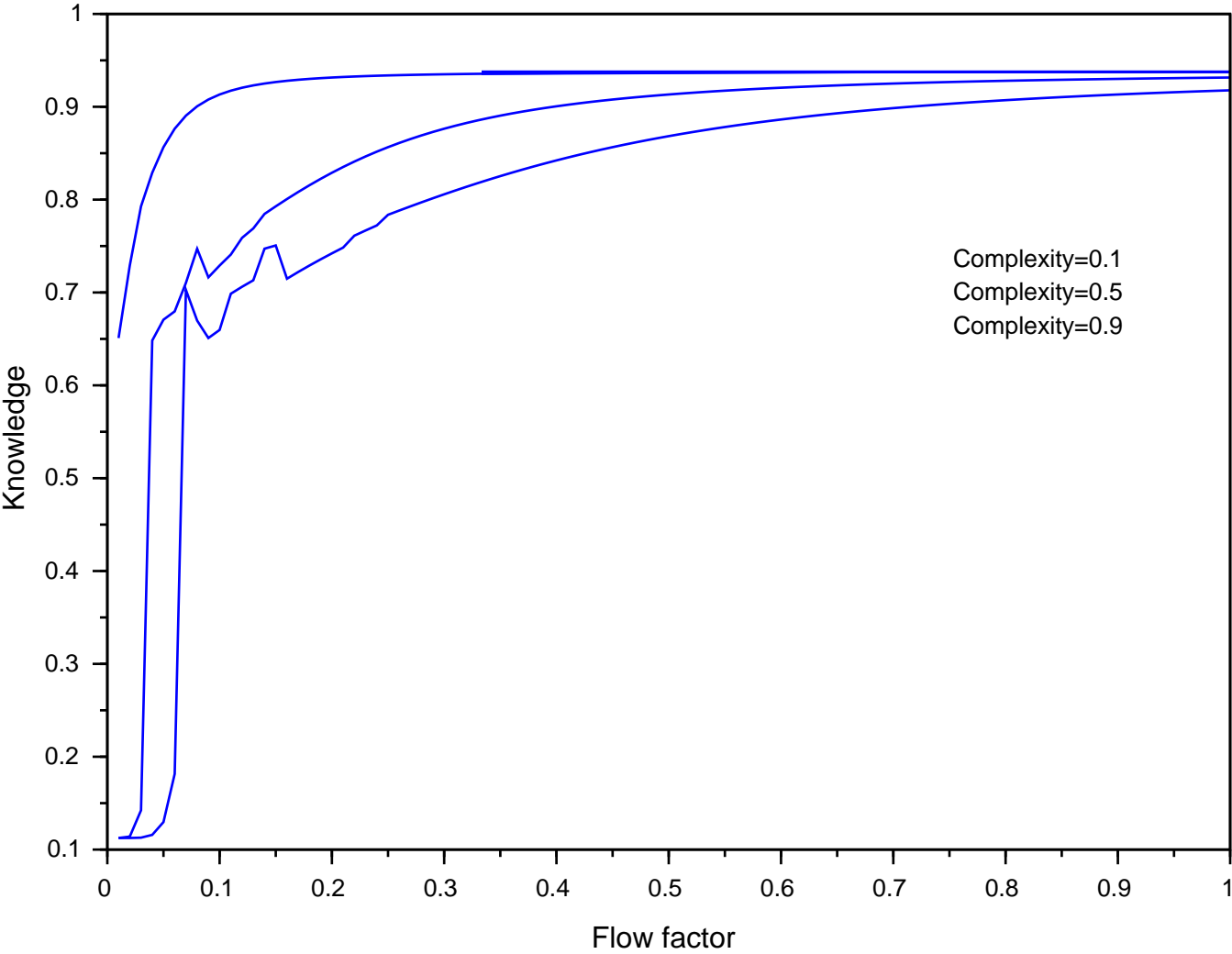
Figure



Figure



Figure



Figure

